# POLYCYSTIC KIDNEY DISEASE GENE HOMOLOGS REQUIRED FOR MALE MATING BEHAVIOR IN NEMATODES AND ASSAYS BASED THEREON

## RELATED APPLICATIONS

This application is a divisional of U.S. application Serial No. 09/479,467, filed January 6, 2000, to Paul W. Sternberg and Maureen M. Barr, and entitled "POLYCYSTIC KIDNEY DISEASE GENE HOMOLOGS REQUIRED FOR MALE MATING BEHAVIOR IN NEMATODES AND ASSAYS BASED THEREON". Benefit of priority under 35 U.S.C. §119(e) to U.S. Provisional Application Serial No. 60/115,127, entitled "*CAENORHABDITIS ELEGANS* STRAINS PERTURBED IN POLYCYSTIN FUNCTION" to Paul W. Sternberg and Maureen M. Barr, filed January 6, 1999, is also claimed herein. The subject matter of each of U.S. Provisional Application Serial No. 60/115,127 and U.S. application Serial No. 09/479,467 is incorporated in its entirety by reference.

## FIELD OF INVENTION

Systems and assays for identification of compounds that can be used to treat polycystic kidney disease (PKD) are provided. Nematode orthologs of genes involved in PKD are identified and associated with mating behaviors. In particular, nematodes, such as *Caenorhabditis elegans*, that express mutant and wild-type orthologs of human genes involved in this disease, are used to study the functions of the proteins encoded by the genes, to screen for other genes involved in the disease, to identify mutations involved in the disease, and to screen for drugs that affect PKD. Hence an animal model is provided that permits study of the etiology of polycystic kidney disease and provides a tool to identify the genes and factors involved in the disease pathway, and to identify compounds that may be used to treat or alter the disease progression, lessen its severity or ameliorate symptoms.

## BACKGROUND

### Polycystic Kidney Diseases

Polycystic kidney diseases (PKD) are a group of disorders characterized by the presence of a large number of fluid-filled cysts throughout grossly enlarged kidneys (Gabow *et al.* (1992) *Diseases of the Kidney*, Schrier *et al.*. eds.). In humans, PKDs can be inherited in autosomal dominant (ADPKD) or autosomal recessive (ARPKD) forms. ADPKD is the more common form and is the most common, dominantly-inherited kidney disease in humans, occurring at a frequency of about 1 in 800. ARPKD occurs at a frequency of about 1 in 10,000.

ADPKD is the most common single-gene disorder leading to kidney failure (see, Emmons *et al.* (1999) *Nature 401*:339-340). Since ADPKD is inherited as an autosomal dominant disorder, children of affected parents have a one in two chance of inheriting the disease. Although the kidney is the most severely affected organ, the disease is systemic and affects the liver, pancreas cardiovascular system and cerebro-vascular system. The major manifestation of the disorder is the progressive cystic dilation of renal tubules (Gabow (1990) *Am. J. Kidney Dis. 16*:403-413), leading to renal failure in half of affected individuals by age 50. Microdissection, histochemical and immunologic studies show that cysts in ARPKD kidneys arise from focal dilations of medullary collecting ducts (McDonald (1991) *Semin. Nephrol. 11*:632-642). Although end-stage renal failure usually supervenes in middle age (ADPKD is sometimes called adult polycystic kidney disease), children may occasionally have severe renal cystic disease.

ADPKD-associated renal cysts may enlarge to contain several liters of fluid and the kidneys usually enlarge progressively causing pain. Other abnormalities such as hematuria, renal and urinary infection, renal tumors, salt and water imbalance and hypertension frequently result from the renal defect. Cystic abnormalities in other organs, including the liver, pancreas, spleen and ovaries are commonly found in ADPKD. Massive

liver enlargement can causes portal hypertension and hepatic failure. Cardiac valve abnormalities and an increased frequency of subarachnoid and other intracranial hemorrhage have also been observed in ADPKD. Progressive renal failure causes death in many ADPKD patients and dialysis and transplantation are frequently required to maintain life in these patients.

Numerous biochemical abnormalities associated with this disease also are observed. These include defects in protein sorting, the distribution of cell membrane markers within renal epithelial cells, extracellular matrix, ion transport, epithelial cell turnover, and epithelial cell proliferation.

Three distinct loci have been shown to cause phenotypically indistinct forms of the AKPKD in humans. These include polycystin-1 (PKD1) on chromosome 16, polycystin-2 (PKD2) on chromosome 4, and polycystin-3 (PKD3) (see, *e.g.*, Reeders *et al.* (1985) *Nature 317*:542-544; Kimberling *et al.*. (1993) *Genomics 18*:467-472; Daoust *et al.* (1995) Genomics, 25:733-736). The ARPKD mutation is on human chromosome 6 (Zerres *et al.* (1993) *Nature Genet. 7*:429-432). Two proteins polycystin-1 (PKD1) and polycystin-2 (PKD2) are defective in human autosomal dominant polycystic kidney disease.

Mutations in either PKD1 or PKD2 cause almost indistinguishable clinical symptoms. Mutations in PKD1 or PKD2 account for 95% of autosomal dominant polycystic disease (Torres *et al.* (1998) Current Opinion in Nephrology and *Hypertension 7*:159-169) with greater than 85-90% of disease incidence being due to mutations in PKD1.

The human PKD1 protein is an approximately 4,300 amino-acid integral-membrane glycoprotein with a large amino-terminal extracellular domain and a small, carboxy-terminal cytoplasmic tail. The human PKD1 gene (see, *e.g.*, U.S. Patent No. 5,891,628), including the complete nucleotide sequence of the gene's coding region (se SEQ ID No. 1) and encoded amino acid sequence, is known (see, SEQ ID No. 2). The

predicted structure of the domains suggested that it is involved in cell-cell interactions or in interactions with the extracellular matrix. The PKD2 protein has similarities to PKD1, but its topology and domain structure suggest that it might act as a subunit of a cation channel. These proteins have been shown to interact directly (Mochizuki *et al.* (1996) *Science 272*:1339-1342, Qian (1997) *Nature Genetics 16*:179-183).

Although these genes have been implicated in the disorders their role in it etiology is not established. In addition, while studies of kidneys from ADPKD patients exhibit a number of different biochemical, structural and physiological abnormalities, the disorder's underlying causative biochemical defect is not known. Hence the molecular mechanisms leading to cyst enlargement and progressive loss of renal function in the PKDs are not understood. Presently there are no cures or effective treatments, other than palliative treatments, for these diseases. Hence there is a need to understand the underlying biochemistry and physiology of the ADPKD and to provide treatments.

Therefore, it is an object herein to provide a means to identify the underlying biochemistry and genetics of these diseases and to provide a means to identify compounds for use in treatment of these diseases.

## SUMMARY

Isolated genes, cDNA and encoded proteins from nematodes that participate in a pathway leading to an observable phenotype are provided. In particular, it is shown herein, that a mutation in *C. elegans*, which gives rise to males that are defective in certain aspects of mating behavior, lies in a gene designed herein *lov-1* (location of vulva), and that this gene is an ortholog of the mammalian, particularly human, PKD1 gene. A mutation in a gene designated *pkd-2* herein also gives rise to these behaviors. This gene is shown to be an ortholog of the mammalian, including human, PKD2 gene.

The expression pattern of *lov-1* and *pkd-2* was studied and it was found that promoter sequences of both genes cause reporter genes to be

expressed in the rays and the hook sensory neurons required for 'response" and vulva location. Thus showing that the LOV-1 and PKD-2 proteins are involved in chemosensory or mechanosensory signal transduction in sensory neurons.

Hence genes that are components of a pathway in nematodes are provided and are shown to be linked to observable behaviors. Each of the encoded proteins, LOV-1 and PKD-2 are components in a pathway, which appears to be a signal transduction pathway, that leads to the observed phenotype. The genes from the nematode *Caenorhabditis elegans* are exemplified herein.

The pathway is shown to be homologous to the pathway in which the human polycystins, PKD1 and PKD2, participate. In particular, it is shown herein, that a mutation in nematodes, which gives rise to males that are defective in mating behavior, lies in a gene designated herein *lov-1* (location of vulva). This gene, *lov-1*, is shown herein to be required for two male sensory behaviors, 'response' and 'location of vulva' (Lov).

A second gene, designated *pkd-2*, that affects this behavior in a similar manner is also identified and provided herein. The encoded proteins are also provided. The gene, cDNA, and encoded protein is also provided. In an exemplary embodiment, the *C. elegans* genome sequence was used to isolate *pkd-2*. This gene is a nematode ortholog of the mammalian, particularly human PKD2 gene. Strains that contain knock-out mutants of this gene also exhibit the defective mating behaviors.

In an exemplary embodiment, provided herein are the *C. elegans* genes, designated *lov-1* and *pkd-2*. SEQ ID No. 3 sets forth the complement (*i.e.*, the non-coding strand) of the *lov-1* gene from *C. elegans*. SEQ ID No. 4 sets forth the sequence of amino acids of the protein (N-terminus to C-terminus)). SEQ ID No. 5 sets forth the complement (*i.e.*, the non-coding strand) of the *C. elegans pkd-2* gene from *C. elegans*. SEQ ID No. 6 sets forth the encoded sequence of amino acids.

Also provided are the mutants of the genes, *lov-1*, and *pkd-2* and the resulting mutant encoded proteins. Nucleic acid molecules encoding mutants of these genes are also provided. For example, deletion mutants of these genes, particularly deletion mutants that substantially or completely knock-out gene product function, are provided. Thus, nucleic acid molecules containing deletions of each of these genes and deletion mutants that alter the phenotype of nematodes, such as *C. elegans*, that contain these mutant genes are also provided. Constructs, vectors, plasmids and strains containing each of the nucleic molecules are also provided. Also provided are strains defective in these genes.

Also provided are strains containing the mutant nucleic acids. Strains that manifest the defective male sensory behaviors are also provided herein. Constructs containing the genes, vectors containing the constructs, cells containing the vectors and transgenic *C. elegans*. Assays that use these strains of *C. elegans* are also provided.

As noted, it is shown herein that these genes are human homologs of the human genes that encode polycystins, proteins polycystin-1 (PKD1) and polycystin-2 (PKD2), which are defective in human autosomal dominant polycystic kidney disease. Hence, the genes and nematode strains provide model systems for studying this pathway, identifying additional components of the pathway, and for use in drug screening assays to identify compounds affect the pathway and/or compounds that serve as leads for development of drugs for treatment of polycystic kidney disease.

Each gene is shown to affect two sensory behaviors in *C. elegans*. One behavior designated "Response" and refers to the response of males to hermaphrodites; and the other behavior, designated "Lov" refers to location of the vulva by the male. Strains that are defective in either or both of these genes are also provided. In particular deletion mutants are provided.

By correlating the phenotypic behaviors with wild-type or defects in these genes, nematodes, such as *C. elegans*, can be used to identify other genes involved in this pathway and also means for direct screening for lead candidate compounds for drugs for treatment of PKD. Identification of additional genes necessary for PKD function can provide additional diagnostic tools for PKD. Hence, provided herein are mutant strains of *C. elegans* and assays that use the strains.

Also provided herein are assays that employ the constructs, vectors, plasmids and strains containing each of the nucleic molecules are also provided. In particular, in one type of assays wild-type nematodes are mutagenized or treated with a test compound, and those that exhibit a change in behavior are identified.

In other types of assays, nematodes that are defective in LOV and/or Response are mutagenized or treated with a compound, and those that exhibit a change in behavior are identified. Test compounds or mutations responsible for the change in behavior are identified. Such compounds are candidates for treatment of PKDs.

Among these methods are those that involved contacting a nematode that exhibits normal mating behavior with a test compound; and selecting compounds that result in altered mating behavior, wherein the altered mating behavior comprises alteration in the behavior involving location of vulva and/or response to contact with the hermaphrodite.

Also provided are methods for identifying genes involved in autosomal dominant polycystic kidney disease (ADPKD). Among these methods are those in involving mutagenizing nematodes that exhibit normal mating behavior; and identifying and selecting nematodes that exhibit altered mating behavior, where the altered mating behavior is manifested as an alteration in location of vulva and/or response to contact with the hermaphrodite. The mutated gene(s) responsible for the alteration in behavior are then identified. Databases or libraries of mammalian genes can be screened to identify homologs of these genes,

which can then serve as therapeutic or diagnostic targets or aid in elucidation of the disease pathology.

Methods for identifying compounds that are candidate therapeutic agents for treatment of autosomal dominant polycystic kidney disease (ADPKD) are provided. Among the methods are those in which normal males are treated with a candidate compound. Compounds that result in changes in mating behaviors or changes in mating efficiencies are selected.

Methods for identifying genes involved in the disease pathway are also provided. Among the methods are those in which normal males are mutagenized. Offspring that exhibit changes in mating behaviors or changes in mating efficiencies are selected and mutated genes are identified and shown to be part of the pathway. Mammalian, particularly human, homologs of the mutated genes are then identified. Such genes are likely to be part of the disease pathway. Such genes can serve as therapeutic targets and disease markers for diagnostic.

Other assays use nematode strains that have mutations in either or both of *lov-1* or *pkd-2*. As described herein, suppressor and enhancer genetics can be used to assign functions to genes, to assign genes to pathways, to identify the key switches in these pathways and to provide a sensitive assay to identify new genes in a pathway and lead compounds that modulate the activity of genes and/or gene products in the pathway.

Assays that identify the role of PKD proteins in sensory function are also provided. Since *lov-1* and *pkd-2* are expressed in CEM neurons, they have activity in other sensory functions, such as finding the mating partner at a distance. Accordingly assays using sexual chemotaxis or kinesis are provided. For example, males that are mutagenized or treated with a test compound are placed on a surface containing males and hermaphrodites, and are then observed to assess whether they can choose between males and hermaphrodites. If the male is defective in

this sensory function, it will not distinguish between males and hermaphrodites.

Assays that use dominant negative forms of PKD in nematodes or in other cells to identify mutations and/or compounds that inhibit PKD function are also provided. Transgenic nematodes that express a version of the LOV-1 or PKD-2 protein that inhibits the activity of LOV-1 and/or PKD-2 as assessed by manifestation of the altered LOV and/or response phenotypic behavior(s) are used in these assays. Transgenic nematodes can be produced by any method known to those of skill in the art, including, but are limited to, injection of the nucleic acid into the embryos or cells of the animal. Transgenic nematodes that contain a dominant negative lov-1 or pkd-2 transgene are contacted with a test compound, and compounds that interfere with a remaining activity of the *LOV-1* or *PKD-2* protein are selected. Alternatively, these transgenic nematodes are mutagenized and mutants that lose a remaining activity are selected and the gene or mutation responsible for the loss or that contributes to the loss is identified.

Assays based on localization and trafficking of LOV-1 and/or PKD-2 within a cell or cells are also provided. These assays can identify regulators and factors necessary for synthesis and transport of *LOV-1* and/or *PKD-2* proteins and employ strains in which LOV-1 and PKD-2 are expressed linked to a detectable label, such as a fluorescent protein. These strains are used to assess the effects of compounds or mutagenesis on the trafficking patterns of *LOV-1* and *PKD-2* and cellular location(s) of the proteins in the animal. Identified mutations can be mapped and the genes identified. If mammalian, particularly human, homologs of these identified genes exist, such genes can serve as therapeutic or diagnostic targets and can aid in elucidation of the disease in mammals, particularly humans.

Assays for identification of transcriptional regulators of expression of *lov-1* and/or *pkd-2* are also provided. These assays screen for loss or

alteration of expression of either gene and use transgenic nematodes with a reporter gene, such as a gene encoding a FP or lacZ or other detectable product, linked to the nucleic acid encoding *lov-1* or *pkd-2*. The animal is mutagenized or treated with a test compound and loss of expression or reduction in expression of either gene is assessed. These assays identify regulators of and factors that affect *lov-1* and *pkd-2* expression. Mammalian, particularly human homologs of these regulators and factors are identified. Such regulators and factors can be therapeutic or diagnostic targets, and/or can aid in developing an understanding of the development and progression of PKD in mammals.

Kits for performing the assays, particularly, the drug screening assays, are also provided. The kits include transgenic or wild-type nematodes or both that express either wild-type or a mutant or a transgenic form of *lov-1* and/or *pkd-2*. The nematodes may be on plates, in wells or in any form suitable for the assays. Kits containing nucleic acid encoding either of the two genes or probes based upon these sequences or reporter gene constructions containing all or portions of either or both genes are also provided. The nucleic acids may be in solution, in lyophilized or other concentrated form, or may be bound to a suitable substrate. The kits can include additional reagents for performing the assays, such reagents include any for performing any of the steps of the methods. The kits include instructions for performing the assays.

**DESCRIPTION OF FIGURES**

Figure 1 depicts male mating behavior of *C. elegans.* The hermaphrodite is larger than the male and her vulva is depicted as a slit on the ventral, posterior third of her body. The male tail is place flush on the hermaphrodite, ventral side down. His spicules are depicted by a line in the tail. The hook is anterior to the spicules, the post cloacal sensilla is posterior. Sequence 1 illustrates wild-type male Lov. Sequence 2 represents hook ablated aberrant Lov behavior (passing and slow search).

Sequence 3 portrays *lov-1(sy552)* mutant behavior (passing and eventually stopping).

Figure 2 depicts the molecular nature of *lov-1*. **a**, Genetic and physical maps of the *lov-1* region on chromosome 2. Genetic markers are shown. Boundaries of a *lov-1* deletion *(mnDf21)* and non-deletion *(eDf21)* are indicated. + designate rescue of *lov-1(sy552)* mutant males. Numbers in parentheses indicate the ratio of rescuing stable lines to total stable lines examined. **b**, *lov-1* gene structure. Exons are boxed. Genefinder predicts two ORFs, ZK945.10 (9 exons) and ZK945.9 (19 exons). RT-PCR reveals *lov-1* corresponds to the combination of ZK945.10 and ZK945.9. The arrow indicates the 1059 bp deletion in *lov-1 (sy582Δ)* **c**, lov-1::GFP (green fluorescent protein) expression constructs, patterns, and phenotypes in wild-type background. **d**, *lov-1* encodes a membrane associated protein with homology to the polycystin and voltage-activated channel families. A schematic representation of LOV-1 is shown to demonstrate domains of the protein. These include the amino terminus that is serine/threonine rich with multiple potential glycosylation sites, an ATP/GTP binding domain (indicated by the asterisks), followed by two polycystin blocks of homology. Block 1 is exclusively homologous to PKD1, while Block 2 shows homology with all polycystins and also the family of voltage activated $CA^{2+}$channels. Block 1 is a conserved domain of unknown function, that also occurs at the N-terminus of most 5-lipoxygenases. Identity (%) and number of identical amino acids (in parentheses) between LOV-1 and a particular polycystin is indicated. Although LOV-1 lacks the carboxy terminal coiled-coil domain of all known polycystins, a coiled-coil is predicted in the middle of LOV-1 using the most stringent criteria for the COILS program (data not shown). Y73F8A.B+A was identified in a Blast search of unpublished sequences available through the Sanger Center and is more similar to PKD2 (30% identity, 48% similarity, 13% gaps over 752 aa) than LOV-1 (25% identity, 44% similarity, 14% gaps over 367 aa).

Figure 3 shows the *lov-1* and *pkd-2* genomic structures, constructs, rescue date and expression patterns; the line above *lov-1* indicates the 1,059 bp deletion in *lov-1(sy582Δ)*; numbers in parentheses indicate the ratio of rescuing stable lines to the number of stable lines examined, DN is dominant negative.

Figure 4 shows that *lov-1::GFP1* and PKD-2::GFP2 are colocalized to cell bodies and dendrites and are specifically expressed in adult male sensory neurons; the spicules, hook structure and posteriomost fan region autofluoresce; Arrows indicate neuronal cell bodies and arrowheads denote dendrites or ciliated endings. **a-c** *lov-1::GFP1*: (a) HOB and ray cell bodies (arrows), HOGB dendridic process (arrowhead); (b) HOB and ray process 5 (arrowheads); (c) Ciliated endings in nose tip from male specific cephalic CEM neurons (cell bodies not shown). **d-f** *pkd-2::GFP2*: (d) ray cell bodies (arrow) and ray process 2 (arrowhead); (e) ray process 5 (arrowhead); (f) male-specific cephalic CEM ciliated endings (arrow) Scale bar corresponds to 20 $\mu$m.

## DETAILED DESCRIPTION

### Definitions

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as is commonly understood by one of skill in the art to which this invention belongs. *Caenorhabditis elegans* nomenclature is well understood by those of skill in this area (see, *e.g.*, *Methods in Cell Biology C. elegans* I, and II, Cold Spring Harbor Press Books, Shakes, Epstein eds).

All patents, patent applications and publications referred anywhere herein, including the background, are, unless noted otherwise, incorporated by reference in their entirety. In the event a definition in this

section is not consistent with definitions elsewhere, the definition set forth in this section will control.

As used herein, nematode is intended to refer generally to the class Nematoda or Nematoidea and includes those animals of a slender cylindrical or thread-like form commonly called roundworms. Among those species, members of the genus *Caenorhabditis* are preferred, but species that can be cultured in the laboratory may be used.

As used herein, the term "mutant," as in "nematode mutant" or "mutant nematode," is intended to refer generally to a nematode which contains an altered genotype, preferably stably altered. The altered genotype results from a mutation not generally found in the genome of the wild-type nematode.

As used herein, a mutant gene, such as a mutant *lov-1* or *pkd-2* gene, refers to a gene that is altered, whereby a nematode with such gene, expresses an altered phenotype compared to a nematode with the wild type gene, such as a the genes set forth in SEQ ID Nos. 3 and 5 (which set forth the non-coding strands). Mutations include point mutations, insertions, deletions, rearrangements and any other change in the gene that results in an altered phenotype. Deletion mutants that eliminate the function of the encoded protein (knock-out mutations) are exemplified herein. Not all mutations necessarily completely destroy the activity of the protein.

As used herein, "normal mating behavior" means that the animal exhibits behavior typical of wild-type nematodes with respect to the location of vulva (Lov) and response to of males to hermaphrodites. Thus a male that exhibits "normal mating behavior" upon encountering a hermaphrodite, ceases forward motion, places his tail flush on the hermaphrodite, commences backing along her body, and turns at her ends until he encounters her vulva and stops. This is the behavior of a *lov-1(+)* male. Mutant males defective in *lov-1* frequently do not respond to contact with the hermaphrodite and continue blindly moving forward.

When response is initiated, *lov-1* mutants back and turn normally but pass the vulva at a high frequency. Thus, they can mate with paralyzed or otherwise slow moving hermaphrodites.

As used herein, a mammalian homolog of a nematode gene refers to a gene that encodes a protein that exhibits identifiable sequence homology and conservation of structure. The degree of sequence homology between a mammalian and nematode protein or gene to be considered homologs, depends upon the gene considered but is typically at least about 30% at the protein level. An ortholog will typically have greater sequence similarity, and conservation of structure and often function. Methods and criteria for identifying mammalian, including human, homologs of nematode genes are known to those of skill in the art and involve a comparison of the sequence and structural features of the encoded protein.

As used herein, a dominant negative mutation is a mutation that encodes a polypeptide that when expressed disrupts that activity of the protein encoded by the wild-type gene (see, Herskowitz (1987) *Nature 329*:219-222). The function of the wild-type gene is blocked, a cloned gene is altered so that it encodes a mutant product that inhibits the wild-type gene product in a cell or organism. As a result, the cell or organism is deficient in the product. The mutation is "dominant" because its phenotype is manifested in the presence of the wild-type gene, and it is "negative" in the sense that it inactivates the wild-type gene function. It is possible to do this because proteins have multiple functional sites.

As used herein, a "library" of nematodes is a collection of a plurality of nematodes, typically more than 10, preferably more than 100. Typically a library will include variety of different nematodes and may include wild-type and mutant nematodes and a sufficient number to achieve the intended purpose for which the library is used..

As used herein, a gene encoding *LOV-1* protein refers to a gene (a sequence of nucleotides including introns, and exons, and optionally

transcriptional regulatory sequences) from any nematode that encodes a protein that performs the same function in the nematode as the LOV-1 protein provided herein. Such protein can be identified using the methods provided herein for identifying it in *C. elegans*, or by isolating cDNA encoding the protein using probes constructed from the nucleic acid provided herein to isolate it using standard methods. Typically the coding sequence of the gene provided herein will hybridize along its length to the coding sequence of a related gene under conditions of at least low stringency, preferably moderate stringency, and likely under conditions of high stringency. Nucleic acid encoding a LOV-1 protein includes any nucleic acid molecule, DNA, cDNA, RNA, that encodes a protein that has substantially the sequence of amino acids set forth in SEQ ID No. 4 and encodes a protein that has the same activity as this protein. Minor sequence variations from species to species and even among a species are considered to be substantially the same sequence. Such nucleic acid will hybridize to the nucleic acid encoding the proteins provided herein under conditions of at least low stringency, preferably moderate stringency and more preferably high stringency.

As used herein, a gene encoding *PKD-2* protein from a nematode is similarly defined, except that it has the substantially the same sequence as the sequence of amino acids set forth in SEQ ID No. 6. Having identified these proteins and functions therefor in *C. elegans* permits similar identification in other nematode species.

As used herein, stringency conditions refer to the washing conditions for removing the non-specific probes and conditions that are equivalent to either high, medium, or low stringency as described below:

1) high stringency: 0.1 x SSPE, 0.1% SDS, 65°C
2) medium stringency: 0.2 x SSPE, 0.1% SDS, 50°C
3) low stringency: 1.0 x SSPE, 0.1% SDS, 50°C.

It is understood that equivalent stringencies may be achieved using alternative buffers, salts and temperatures.

As used herein, percentage or amount or degree of sequence identity is used interchangeable with homology and refers to sequence identity or homology determined using standard alignment programs with gap penalties and other parameters set to the manufacturer's default settings. It is understood that for relatively high levels of sequence identity or homology, the particular program selected and/or defaults set for various parameters, do not substantially affect the results. Hence, for example, a requirement for 90% sequence identity of a nucleic acid sequence with another can be determined using any program known to the skilled artisan or manually, and that such percentage can encompass about 85% to 95% identity.

As used herein, reference to a drug refers to a chemical entity, whether in the solid, liquid, or gaseous phase that is capable of providing a desired therapeutic effect when administered to a subject. The term "drug" should be read to include synthetic compounds, natural products and macromolecular entities such as polypeptides, polynucleotides, or lipids and also small molecules, including, but are not limited to, neurotransmitters, ligands, hormones and elemental compounds. The term "drug" is meant to refer to that compound whether it is in a crude mixture or purified and isolated.

As used herein, heterologous or foreign DNA and RNA are used interchangeably and refer to DNA or RNA that does not occur naturally as part of the genome in which it is present or which is found in a location or locations in the genome that differ from that in which it occurs in nature. Heterologous nucleic acid is generally not endogenous to the cell into which it is introduced, but has been obtained from another cell or prepared synthetically. Generally, although not necessarily, such nucleic acid encodes RNA and proteins that are not normally produced by the cell in which it is expressed. Any DNA or RNA that one of skill in the art would recognize or consider as heterologous or foreign to the cell in which it is expressed is herein encompassed by heterologous DNA.

Examples of heterologous DNA include, but are not limited to, DNA that encodes exogenous invertase. Heterologous DNA and RNA may also encode RNA or proteins that mediate or alter expression of endogenous DNA by affecting transcription, translation, or other regulatable biochemical processes.

As used herein, operative linkage of heterologous DNA to regulatory and effector sequences of nucleotides, such as promoters, enhancers, transcriptional and translational stop sites, and other signal sequences refers to the relationship between such DNA and such sequences of nucleotides. For example, operative linkage of heterologous DNA to a promoter refers to the physical relationship between the DNA and the promoter such that the transcription of such DNA is initiated from the promoter by an RNA polymerase that specifically recognizes, binds to and transcribes the DNA in reading frame.

As used herein, a gene containing a heterologous transcriptional or translational or processing control region(s) refers to a nucleic acid molecule or construct that includes coding portion of a gene operatively linked to a such region derived from a different gene. A homologous transcriptional or translational or processing control region(s) refers to a nucleic acid molecule or construct that includes coding portion of a gene operatively linked to a such region derived from the same gene.

As used herein, a promoter region refers to the portion of DNA of a gene that controls expression of DNA to which it is operatively linked. The promoter region includes specific sequences of DNA that are sufficient for RNA polymerase recognition, binding and transcription initiation. This portion of the promoter region is referred to as the promoter. In addition, the promoter region includes sequences that modulate this recognition, binding and transcription initiation activity of the RNA polymerase. These sequences may be cis acting or may be responsive to trans acting factors. Promoters, depending upon the nature of the regulation, may be constitutive or regulated. A constitutive

promoter is always turned on. A regulatable promoter requires specific signals to be turned on or off. A developmentally regulated promoter is one that is turned on or off as a function of development.

As used herein, regulatory sequences include, sequences of nucleotides that function, for example as transcriptional and translational control sequences. Transcriptional control sequences include the promoter and other regulatory regions, such as enhancer sequences, that modulate the activity of the promoter, or control sequences that modulate the activity or efficiency of the RNA polymerase that recognizes the promoter, or control sequences are recognized by effector molecules, including those that are specifically induced by interaction of an extracellular signal with a cell surface protein. For example, modulation of the activity of the promoter may be effected by altering the RNA polymerase binding to the promoter region, or, alternatively, by interfering with initiation of transcription or elongation of the mRNA. Such sequences are herein collectively referred to as transcriptional control elements or sequences. In addition, transcriptional controls sequences, include sequences of nucleotides that alter translation of the resulting mRNA, thereby altering the amount of a gene product.

As used herein, a reporter gene refers to a gene that encodes a detectable product. Such genes are well known to those of skill in the art and include, but are not limited to, genes encoding fluorescent proteins, particularly the well-known green fluorescent proteins, *lacZ*, enzymes and other such reporters known to be expressible and detectable in nematodes. These genes are linked to a gene of interest whereby upon expression a detectable fusion protein is produced. For purposes herein, such fusions are exemplified using an aequorin GFP (see, Chalfie *et al.* (1994) *Science 263*:802-805; see, also U.S. Patent No. 5,741,668), but any such protein may be used. For example, GFP from *Aequorea victoria* contains 238 amino acids, absorbs blue light and emits green light; it has

been cloned and its sequence characterized; various mutants are also well known. Nematode optimized codons may be selected.

As used herein, a reporter gene construct is a nucleic acid molecule that includes a reporter gene operatively linked to transcriptional control sequences. Typically the construct will also include all or a portion of a the gene of interest, which herein is *lov-1* and/or *pkd-2*, and the reporter gene will be under the control of the *lov-1* or *pkd-2* promoter and other regulatory regions. By operatively linked is meant linked whereby an in-frame fusion protein is produced upon expression of the construct and whereby the reporter gene product is active (*i.e.* produces a detectable signal or is active). The reporter gene may be linked to the 3' or 5' end or in any other orientation whereby it is expressed and operates as a reporter.

As used herein, isolated, substantially pure DNA refers to DNA molecules or fragments purified according to standard techniques employed by those skilled in the art, such as those described in Sambrook *et al.* (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY).

As used herein, expression refers to the process by which nucleic acid is transcribed into mRNA and translated into peptides, polypeptides, or proteins. If the nucleic acid is derived from genomic DNA, expression may, if an appropriate eukaryotic host cell or organism is selected, include splicing of the mRNA.

As used herein, cloning vehicle or vector, which are used interchangeably, refers to a plasmid or phage DNA or other DNA molecules that replicate autonomously in a host cell, and that include one or a small number of endonuclease recognition sites at which such DNA may be cut in a determinable fashion without loss of an essential biological function of the vehicle, and into which DNA may be spliced in order to bring about its replication and cloning. The cloning vehicle may further contain a marker suitable for use in the identification of cells

transformed with the cloning vehicle. Markers, include but are not limited to, tetracycline resistance and ampicillin resistance.

Appropriate expression vectors are well known to those of skill in the art and include those that are replicable in eukaryotic cells and/or prokaryotic cells. Such expression vectors may remain episomal or may integrate into the host cell genome. Expression vectors suitable for introducing heterologous DNA into plants and into host cells in culture, such as mammalian cells and methylotrophic yeast host cells, are known to those of skill in the art. It should be noted that, because the functions of plasmids, vectors and expression vectors overlap, those of skill in the art use these terms, plasmid, vector, and expression vector, interchangeably. Those of skill in the art, however, recognize what is intended from the purpose for which the vector, plasmid or expression vector is used.

As used herein, integrated into the genome means integrated into a chromosome or chromosomes.

As used herein, a "fragment" of a protein refers to any portion of a protein that contains less than the complete amino acid sequence of the protein but that retains a biological or chemical function of interest.

As used herein, expression vector or expression vehicle refers to such vehicle or vector that capable, after transformation into a host, of expressing a gene cloned therein. The cloned gene is usually placed under the control of (i.e., operably linked to) certain control sequences such as promoter sequences. Expression control sequences will vary depending on whether the vector is designed to express the operably linked gene in a procaryotic or eukaryotic host and may additionally contain transcriptional elements such as enhancer elements, termination sequences, tissue-specificity elements, and/or translational initiation and termination sites.

As used herein, a variant of a protein refers to a protein substantially similar in structure and biological activity to

either the entire protein or a fragment thereof. Thus, provided that two proteins possess a similar activity, they are considered variants as that term is used herein even if the composition or secondary, tertiary, or quaternary structure of one of the molecules is not identical to that found in the other, or if the sequence of amino acid residues is not identical.

It is also understood that any of the proteins or portions disclosed herein may be modified by making conservative amino acid substitutions and the resulting modified subunits are contemplated herein. Suitable conservative substitutions of amino acids are known to those of skill in this art and may be made generally without altering the biological activity of the resulting molecule. Those of skill in this art recognize that, in general, single amino acid substitutions in non-essential regions of a polypeptide do not substantially alter biological activity (see, e.g., Watson et al. *Molecular Biology of the Gene*, 4th Edition, 1987, The Benjamin/Cummings Pub. Co., p.224). Such substitutions are preferably, although not exclusively, made in accordance with those set forth in TABLE 1 as follows:

**TABLE 1**

| Original residue | Conservative substitution |
|---|---|
| Ala (A) | Gly; Ser |
| Arg (R) | Lys |
| Asn (N) | Gln; His |
| Cys (C) | Ser |
| Gln (Q) | Asn |
| Glu (E) | Asp |
| Gly (G) | Ala; Pro |
| His (H) | Asn; Gln |
| Ile (I) | Leu; Val |
| Leu (L) | Ile; Val |
| Lys (K) | Arg; Gln; Glu |
| Met (M) | Leu; Tyr; Ile |
| Phe (F) | Met; Leu; Tyr |
| Ser (S) | Thr |
| Thr (T) | Ser |
| Trp (W) | Tyr |
| Tyr (Y) | Trp; Phe |
| Val (V) | Ile; Leu |

Comparable mutations may be made at the nucleotide sequence level.

Other substitutions are also permissible and may be determined empirically or in accord with known conservative substitutions. Any such modification of the polypeptide may be effected by any means known to those of skill in this art. Mutation may be effected by any method known to those of skill in the art, such as by chemicals or radiation, and also including site-specific or site-directed mutagenesis of DNA encoding the protein and the use of DNA amplification methods using primers to introduce and amplify alterations in the DNA template.

As understood by those skilled in the art, assay methods for identifying compounds, such as antagonists and agonists, that modulate functioning of a protein or protein or pathway, generally require comparison to a control. One type of a "control" system is one that is treated substantially the same as the system, such as a worm, exposed to the test compound except that the control is not exposed to the test compound. Another type of a control may be one that is identical to the test system, except that it does not express the gene or protein of interest. In this situation, the response of a test system is compared to the response (or lack of response) of the control to the test compound, when each cell is exposed to substantially the same reaction conditions in the presence of the compound being assayed.

As used herein, treatment means any manner in which the symptoms of a condition, disorder or disease are ameliorated or otherwise beneficially altered.

As used herein, amelioration of the symptoms of a particular disorder by administration of a particular pharmaceutical composition refers to any lessening, whether permanent or temporary, lasting or transient that can be attributed to or associated with administration of the composition.

As used herein, a composition refers to any mixture of two or more components. It may be solution, suspension, or any other mixture.

As used herein, biological activity refers to the in vivo activities of a compound or physiological responses that result upon in vivo administration of a compound, composition or other mixture. Biological activity, thus, encompasses therapeutic effects and pharmaceutical activity of such compounds, compositions and mixtures.

**Nematodes as disease models**

Nematodes serve as model organisms for the study of gene expression. *Caenorhabditis elegans* is representative of nematodes. It is a small, freeliving bacteriovorous soil nematode that is a member of the *Rhabditidae*, a large and diverse group of nematodes found in terrestrial habitats. Some rhabditids are pathogenic to or parasitic on animals. In common with other nematodes, *C. elegans* develops through four larval stages (also called juveniles) that are separated by moults. The lifecycle takes about 3 days at 20 ° C.

*C. elegans* is only 1 mm long and can be handled in a manner similar to microorganisms, including growth on petri plates seeded with bacteria. In the laboratory, *C. elegans* is fed on *E. coli*. It has a transparent body and all somatic cells ( 959 female; 1031 male) are visible with a microscope.

Although it is a primitive organism, it shares many of the essential biological characteristics, including embryogenesis, morphogenesis, development and aging that are central problems of human biology. The worm is conceived as a single cell that undergoes a complex process of development, starting with embryonic cleavage, proceeding through morphogenesis and growth to the adult. It has a nervous system with a 'brain' (the circumpharyngeal nerve ring), It exhibits definable behaviors, and is capable of rudimentary learning. It produces sperm and eggs, mates and reproduces. After reproduction it gradually ages, loses vigor and dies. Its average life span is 2-3 weeks.

Adult *C. elegans* are usually self-fertilizing protandrous hermaphrodites. As a result homozygous mutant stocks can be readily

generated. The hermaphrodite gonad first produces germ cells that differentiate as sperm (about 250 sperm are produced) and then produces eggs. The fecundity is determined by the sperm supply.

Nematodes, particularly *C. elegans*, is one of the most thoroughly understood of all multicellular organisms. The biology of its nervous system, which contains 302 neurons, is well-documented. Many *C. elegans* genes used have counterparts in mammals, including humans. At least half of the C. elegans genes and proteins that have been characterized have structures and functions similar to mammalian genes. These include genes encode enzymes, proteins necessary for cell structure, cell surface receptors and genetic regulatory molecules.

Animals from man to worm have most of their protein families in common and humans frequently have four to five close analogs of a protein family member, where worms have only one. Essentially all genes and pathways shown to be important in cell-, developmental- and disease-biology have been found to be conserved between worm and human. This conservation applies to the number and type of protein families, gene structure, the hierarchy of genes in genetic pathways and even gene regulation.

A consequence of this conservation is that human genes can be inserted into the worm genome, to functionally replace the worm genes even in complex cell biological and signal transduction pathways. Conversely, key worm genes identified using genetics can be used to trigger specific biochemical processes in human cells and to serve as models for the human genes.

**Genetics Nomenclature**

*C. elegans* is diploid and has five pairs of autosomal chromosomes (designated I, II, III, IV and V) and a pair of sex chromosomes (X) that determine gender. XX is a hermaphrodite and XO is male. Males are found rarely (about 0.05% of normal lab populations). The commonest lab strain, and the designated "wild-type" strain, is called N2.

For historical reasons *C. elegans* nomenclature is different from other species. Loci have a 3-letter dash one number designation. The letters are an acronym for the phenotype and the number is consecutive. Alleles have a single or double letter followed by a number. The letter identifies the isolating laboratory. Strains have a letter(s) number designation. The letters identify the isolating laboratory (i.e. AB100 abc-1(xy1000) Strain AB100 which carries the xy1000 allele of abc-1. The chromosomal location can be added: AB100 abc-1(xy1000) I. Multiple mutant alleles carried in one strain are organized by chromosome, and chromosomes separated by semicolons. Heterozygous nematodes are designated by a abc-1/+ notation. Hence abc-1(+) indicates the wild-type (N2 strain) copy of the gene. Proteins are capitalised and not italicized. ABC is the protein product of abc-1.

Rearrangements, duplications and deficiencies have a letter prefix (indicating the isolating lab) a Dp (pronounced dupe, for duplication) or Df (pronounced dif for deficiency) and a number (*i.e.*, xyDp1 is duplication number 1 from xy and xyDf1 is deficiency number 1 from xy lab). Transgenic strains carrying the transgene as a free extrachromosomal array are designated as follows: xyEx1[abc-1(+)] is a transgenic strain carrying the wt copy of abc-1.

**The C. elegans Genome**

The *C. elegans* genome, which is 97 Mb, contains six approximately equally sized chromosomes (5 autosomes, one X) and it has been sequenced (see,(1998) *Science 282*:2012-2018) and is publicly available. The 97 Mb encodes a predicted 19,099 protein-encoding genes; although as shown herein, there remain ambiguities. Over 60,000 cDNA fragments have been tag sequenced and 101000 ESTs deposited. These "expressed sequence tags" or ESTs offer a set of snapshots of gene expression in the nematode, and have identified around half of the organism's genes. The cDNA data is used in the prediction of genes from the genome sequence along with database

searches for similarities between *C. elegans* genes and those of other organisms such as humans. This estimate is based on the correspondence between genomic DNA sequence and cDNA sequences, and on the prediction of coding genes from genomic sequence.The genome data (and much else besides) is collated into an available database ACeDB, written for the *C. elegans* project. A physical map of the genome, which is publically available in the *C. elegans* genome database ACeDB, has been constructed. The map is based on 17,000 cosmid clones of genomic DNA (insert size 35-40 kb). These clones were "fingerprinted" using restriction enzymes, and the fingerprints used to order the clones in overlapping contiguous sets, or contigs. These cosmid contigs have been supplemented by a set of 3,000 yeast artificial chromosome clones (insert sizes 100 kb and above). Because the yeast host tolerates sequences that *E. coli* does not, the YAC clones can "bridge" gaps between contigs of cosmids. With these two resources, contigs covering >95% of all the chromosomes have been assembled. The clones are freely available for researchers, and the 3,000 YAC clones are available as an array on a filtermat, arranged in approximate chromosomal order, for screening purposes.

The genomes of other nematodes are in the same size range. *Brugia malayi*, a filarial parasite of humans, has a genome of 100 Mb; *Ascaris suum*, the pig roundworm, has a larger germ line genome which undergoes somatic diminution.

**Identification of the genes associated with the location of vulva and response behaviors**

**The behaviors**

The six sub-steps of the stereotyped copulatory sequence has been correlated with the function of individual neurons, and behavioral mutants have been isolated (Liu *et al. Neuron 14*:79-89). *C. elegans* male mating behavior includes a series of steps: response to contact with the hermaphrodite, backing along the body of the hermaphrodite, turning

around her head or tail, location of the vulva, insertion of the two copulatory spicules into the vulva and sperm transfer. Sensory structures and neurons that participate in each of these steps have been identified: the sensory rays mediate response to contact and turning; the hook, the postcloacal sensilla and the spicules mediate vulva location; and the spicules also mediate spicule insertion and regulate sperm transfer.

Thus, the stereotyped mating behavior of the *Caenorhabditis elegans* male comprises several substeps: response backing, turning, vulva location, spicule insertion, and sperm transfer (Fig. 1). The complexity of male mating behavior is reflected in the sexually dimorphic anatomy and nervous systems of the male and hermaphrodite (Hodgkin, J. (1988) in *The Nematode C. elegans* (ed. Wood, B.) pp. 243-279 (Cold Spring Harbor Laboratory Press, New York). Behavioral functions have been assigned to most male-specific sensory neurons via cell ablations (Liu *et al. Neuron 14*:79-89). Although the hermaphrodite is behaviorally passive, her vulva provides sensory cues to the male.

Vulva location behavior is complex. The male stops and precisely positions his tail over the vulva, coordinates his movement to the hermaphrodite's, and ultimately insert his spicules into the vulva slit and transfers sperm into the uterus. The hook sensory neurons, HOA and HOB, are specifically required for location of vulva (Lov) behavior. Ablation of either HOA or HOB results in a Lov defect whereby the ablated male circles the hermaphrodite without stopping at the vulva (Fig. 1). Eventually, the ablated male begins an alternative search by backing slowly and prodding randomly with his spicules until the vulva is located. The postcloacal sensilla are required for slow search behavior. Vulva location behavior is executed by a minimum of eight sensory neurons with overlapping and redundant functions (Liu *et al. Neuron 14*:79-89).

A genetic analysis of vulva location behavior to investigate how genes specify sensory behavior, beginning with sensory reception was

performed. The mating behavior of existing mutants defective in sensory behaviors including chemotaxis to soluble and volatile odorants, mechanosensation, and osmotic avoidance was first examined. From this survey, it was found that only males with severe defects in all sensory neuron cilia (*osm-4, osm-5, osm-6*, and *che-3*) were Lov defective (Table 2). For example, *osm-6(p811)* males locate the vulva with an efficiency of 32% versus 96% of wild-type (Table 2). These males are also response defective, but not so severely as to prevent observation of the Lov phenotype. The only ciliated cells in *C. elegans* are chemosensory and mechanosensory neurons (White *et al.* (1986) *Philos. Trans. R. Soc. Lond. B Biol. Sci. 314*:1-340). The male tail possesses thirty predicted ciliated sensory neurons (Sulston *et al.* (1980) *Dev. Biol. 78*:542-576), consistent with the observation that ciliated neurons modulate response and Lov. *osm-6::gfp* is expressed exclusively in ciliated neurons, with male-specific expression in four CEM head neurons and neurons of the rays and copulatory spicules (Collet *et al.* (1998) *Genetics 148*:187-200). More detailed examination revealed that *osm-6::gfp* expression begins at the L4 stage in neuronal cell bodies and extends to dendrites as neuronal outgrowth proceeds (data not shown). The RnA and RnB neurons of each ray (ray 1 through ray 9), the HOA and HOB hook neurons, the spicule neurons SPV and SPD, and the PCB postcloacal sensilla neurons accumulate GFP. The *osm-6* expression pattern and mutant phenotypes indicate that OSM-6 might be required for the structure and function of ciliated neurons in the adult male tail. In the hermaphrodite, *osm-6* function is required for nose touch (Kaplan *et al.* (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*:2227-2231), osmotic avoidance, chemotaxis, dye-filling of sensory neurons, thermotaxis, dauer formation, and proper assembly of ciliated sensory endings (Perkins *et al.* (1986) *Dev. Biol. 117*:456-487). Hence, ciliated endings are important for all known sensory behaviors, including Lov.

**TABLE 2. Vulva location behavior of wild-type and mutant males**

| Genotype | vulva location efficiency % | Significantly different from wild-type (p value) | | †n |
|---|---|---|---|---|
| *him-5* (wild-type) | 96 | — | — | 101 |
| *osm-1(e1803)* | 65 | No | (0.0738) | |
| *osm-4(p821)* | 48 | Yes | (0.0004) | |
| *osm-5(p813); him-5* | 26 | Yes | (0.0002) | |
| *osm-6(p811)* | 32 | Yes | (0.0003) | |
| *che-3(e1124)* | 69 | Yes | (0.02666) | |
| *lov-1(sy582Δ)* | 11 | Yes | (<0.0001) | |
| *lov-1(sy582); him-5* | 30 | Yes | (<0.0001) | |

Table 2. *lov-1(sy522); him-5(e1490), lov-1(sy582Δ),* and all cilia defective mutant were also response defective. Males that eventually responded were scored for Lov behavior. †**n** represents the number of males observed, each for a minimum of 10 vulva encounters per male. Mann-Whitney tests determined p values. The following non-cilia-defective osmotic avoidance (*osm*), mechanosensory defective (*mec*), chemosensory defective (*che*), odorant response abnormal (*odr*) and dauer formation defective (*daf*) mutants were also examined and found to be normal for response and Lov behavior: *osm-3(e1806); him-5(e1490), osm-7(n1515), osm-8(n1518), osm-10(n1604),osm-11(n1604), osm-12(n1606), mec-3(e1338) him-8(e1489), mec-4(e1611), mec-5(e1340), mec-7(n434), mec-7(e1343), mec-8(e398), mec-9(e1494), che-112, odr-1(n1936), odr-2(n2145), odr-3(n2150), odr-4(n2144ts), odr-5, odr-6(kyl), odr-7(ky4, odr-10(ky32)* and *daf-11(m47ts).*

Provided herein are mutants that are defective in location of the vulva (Lov). Lov mutant males are unable to execute this step. In addition, these males are also defective in the first sub-step, 'response'. Response and vulva location depend on two types of male sensory structure: the first is a set of nine pairs of rays, which project out of the tail on each side; and the second is a hardened cuticular structure called the hook, which contains two sensory neurons. These mutants were used to identify the genes involved in these behaviors.

**Identification and cloning of the lov-1 gene**

To elucidate the molecular basis of behavior and sensory the mutants are studied and genes associated with the behaviors are identified. A gene designated *lov-1* that is required for two male sensory behaviors, response and location of vulva (Lov) is described herein. It is also associated with other sensory behaviors controlled by the CEM neurons.

This gene, *lov-1*, encodes a putative membrane protein with a mucin-like, serine-threonine rich amino terminus (Carraway *et al.* (1995) *Trends Glycoscience Glycotechnology 7*:31-44) followed by two blocks of homology to human polycystins encoded by the autosomal dominant polycystic kidney disease (ADPKD) genes (Torres *et al.* (1998) Current Opinion in Nephrology and *Hypertension 7*:159-169). LOV-1 and human PKD1 are 26% identical in block 1. Block 2 also shows 20% identity between LOV-1, all identified polycystins (PKD1, PKD2, and PKDL), and the family of voltage-activated channels (Torres *et al.* (1998) Current Opinion in Nephrology and *Hypertension 7*:159-169). Overall, LOV-1 is the closest *C. elegans* homolog of PKD1. The polycystin/channel domain (block 2) of LOV-1 is required for function. *Lov-1* is specially expressed in adult male sensory neurons of the rays, hook, and head, mediating response, Lov, and potentially chemotaxis to hermaphrodites, respectively (Liu *et al. Neuron 14*:79-89, Ward *et al.* (1975) *J. Comp. Neurol. 160*:313-337). Localization of *lov-1* to neuronal cell bodies and ciliated sensory endings is consistent with a role in either chemo- and/or mechanosensory reception and signaling. Human PKD proteins might similarly be involved in sensory reception during osmoregulation, organogenesis and/or organ maintenance.

**Cloned genes and encoded proteins**

To identify genes specifically required for male sensory behaviors, mutants defective in Lov were screened. *Lov-1(sy552)* males have specific response and Lov defects. Upon encountering a hermaphrodite, a *lov-1(+)* male ceases forward motion, places his tail flush on the hermaphrodite, commences backing along her body, and turns at her ends until he encounters her vulva and stops. Mutant males defective in *lov-1* frequently do not respond to contact with the hermaphrodite and continue blindly moving forward. When response is initiated, *lov-1* mutants back and turn normally but pass the vulva at a high frequency. The response and vulva location ability of *lov-1(sy552)* is 30% that of *lov-1(+)* males

(Table 2). Spicule insertion and sperm transfer behaviors are unaffected. *lov-1(sy552)* males exhibit high mating efficiency with severely paralyzed *unc-52* hermaphrodites but sire few progeny with actively moving *dpy-17* hermaphrodites. Differences between mating efficiencies is partner-dependent. A paralyzed partner is an easier target for the *lov-1* mutant male who is defective in response and Lov but unimpaired in the behaviors of backing, turning, spicule insertion, and sperm transfer. The behavioral defects of *sy552* are limited to male mating. *Lov-1(sy552)* mutants appear normal for other sensory behaviors including egg laying, nose touch, tap, mechanosensation, and osmotic avoidance.

The *lov-1* gene was cloned by genetic mapping and transformation rescue of the *sy552* behavioral defects (Fig. 2a). *mnDf2l/sy552, mnDf83/sy552* and *sy552/sy552* males are phenotypically indistinguishable; therefore, *sy522* is reduction or loss of function mutation in *lov-1*. This conclusion is supported by the observed recessive nature of *sy552*. A 16.9 kb HindIII subclone (plov-1.1) of the cosmid ZK945 rescued response and Lov defects of *sy552* (Fig. 2a). Both a 6.7 kb HindIII-BamHI fragment from plov-1.1 (plov-1::GFP1) and a 14.1 kb HindIII-StuI frameshift in plov-1.1 (plov-1.3) fail to rescue *sy552* defects (Fig. 2b) yet act in a dominant negative (DN) manner in wild-type males with respect to Lov behavior (Fig. 2c). Wild-type males expressing either plov-1::GFP or plov-1.3 are Lov defective. These transgenic males exhibit a wild-type response to hermaphrodite contact. Without being bound by a theory, the differences in *sy552* and transgenic DN phenotypes might be attributed to dosage or mosaicism.

Figure 2b illustrates the intron-exon boundaries of the *lov-1* gene. Using RT-PCR with *lov-1* specific primers and *him*-5 mRNA, it was found that *lov-1* encodes one transcript corresponding to Genefinder-predicted ORFs, ZK945.10 and ZK945.9 (Fig. 2b), which had been thought to be two genes. *Lov-1* encodes a predicted 3178 amino acid membrane-bound protein (see SEQ ID Nos. 3 and 4) with a serine-threonine rich

extracellular domain homologous to mucins (Carraway *et al.* (1995) *Trends Glycoscience Glycotechnology* 7:31-44), a polycystin homology block 1 (26% identity), and a carboxy terminal polycystin block 2 with 20% identity to polycystin proteins 1, 2, and 2, encoded by the PKD1, PKD2, and PKDL (polycystic kidney disease) genes, respectively (Fig. 2d). A Kyte-Doolittle hydropathy plot predicts multiple transmembrane domains; although no signal peptide is predicted in LOV-1. Mucins are highly glycosylated extracellular proteins thought to serve cell adhesion and/or protective functions (Carraway *et al.* (1995) *Trends Glycoscience Glycotechnology* 7:31-44).

Similarity between exons W (for PKD1 only), X, Y, Z, AA, BB, and CC of *lov-1* and PKD1, PKD2, and the family of voltage-activated calcium and potassium channels in the six transmembrane spanning region has been observed (Mochizuki *et al.* (1996) *Science 272*:1339-1342). This extends to PKDL (Nomura *et al.* (1998) *J. Biol. Chem. 273*:25967-25973). LOV-1 lacks the $Ca^{2+}$ binding EF-hand of polycystin 2 and L, and a coiled-coil domain of all three polycystins (Fig. 2d), which has been shown to mediate hetero- and homotypic interactions between polycystin 1 and polycystin 2 (Qian (1997) *Nature Genetics 16*:179-183; Tsiokas *et al.* (1997) *Proc. Natl. Acad. Sci. USA 94*:6965-6970). Block 2 also shows limited homology with the trp (transient receptor potential) family of channels (Montell *et al.* (1989) *Neuron 2*:1313-1323). The critical difference between voltage-gated and trp channels is the presence of a positively charged S4 transmembrane domain that acts as a voltage sensor (Montell *et al.* (1989) *Neuron 2*:1313-1323). LOV-1 more closely resembles voltage-gated channels in this respect. A frameshift disruption in *lov-1* (plov-1.3) one residue away from a corresponding nonsense mutation in human PKD2 (Mochizuki *et al.* (1996) *Science 272*:1339-1342) destroys the ability to rescue *lov-1(sy552)*, as mentioned above. The construct plov-1.3 encodes a truncated protein lacking the polycystin block 2/channel domain. These results demonstrate that the polycystin

block 2/channel domain is essential for LOV-1 function, and indicate that functional as well as structural similarities might exist between LOV-1 and PKD-2. LOV-1 also possesses a nucleotide-binding domain (Fig. 2d) that is not present in the human polycystins. The structure of LOV-1 is also indicative of a role in signal transduction.

The *lov-1* gene product appears to be a membrane spanning protein that includes an extracellular domain with a serine/threonine-rich mucin-like domain, an ATP-binding domain, and small cytoplasmic tails that mediate interaction with other members of the pathway, including a *pkd-2* gene product that is also a membrane spanning protein, with six membrane domains, and a cytoplasmic EF-hand. Interaction of these proteins lead to the observed phenotypic response. In *c. elegans* this response can be detected as a clearly identifiable phenotype. Hence, *c. elegans* and mutants thereof can serve as a test system for identifying compounds that alter this pathway and also for identifying other gene products involved in the pathway.

***lov-1* gene**

In an exemplary embodiment, the complement of the nucleic acid sequence of the *lov-1* gene from *C. elegans* is provided. Corresponding genes from other nematodes may be identified, such as by using the nucleic acid provided herein and screening an appropriate library, genomic or cDNA library, using standard procedures. Alternatively, databases of sequence may be searched and the genes from other nematodes homologous to those provided herein identified, again using standard searching and alignment programs.

SEQ ID NO. 3 is the complement of the genomic sequence of the *lov-1* gene. It includes open reading frames (ORFs) between nucleotides 15760 to 27880 of cosmid ZK945 (nucleotides 1 to 12121 of SEQ ID NO.3) and nucleotides 1-564 of cosmid F27E5 (nucleotides 12122 to 12685 of SEQ ID NO.3). It was found herein, however, that ZK945 and F27E5 overlap from nucleotides 27881 to 27981 and nucleotides 1 to

101, respectively (the overlap region includes nucleotides 12122 to 12222 in SEQ ID NO.3), thereby providing a single, rather than two, ORFs.

It been thought that the open reading frame in cosmid ZK945 (the "ZK945.9" gene; nucleotides 1 to 9164 of SEQ ID NO.3), and the open reading from in cosmid F27E5 (the "ZK945.10" gene; nucleotides 9415 to 12685 of SEQ ID NO.3) encoded two genes. DNA sequence analysis of RT-PCR generated cDNA clones from *him-5(e1490)* RNA revealed three exons (**exons I, J and K** in Figure 2B) in the junction between ZK945.10 and ZK945.9: one from nucleotides 25195 to 25742 of the ZK945 cosmid (nucleotides 9436 to 9983 of SEQ ID NO. 3); a second from nucleotides 25071 to 25151 of the ZK945 cosmid (nucleotides 9312 to 9392 of SEQ ID NO. 3); and a third initiating at position 25021 in the ZK945 cosmid (nucleotide 9262 of SEQ ID NO. 3). This demonstrated that the *lov-1* gene encodes one large transcript corresponding to ORFs in ZK945.10 and ZK945.9, spanning what had previously been thought to encode two proteins.

As noted above, Figure 2B depicts the *lov-1* genomic structure (exons shown as boxes, introns as lines). With reference to Figure 2B, the coding sequence in the gene set forth in SEQ ID No. 3 (noting that SEQ ID 3 sets forth the non-coding strand) is as follows:

Complement (Join (12500...12685) - Exon A; (12266...12451) - Exon B; (12085...12217) - Exon C; (11683...11823) - Exon D; (11498...11637) - Exon E; (11128...11452) - Exon F; (10268...10899) - Exon G; (10138...10216) - Exon H; (9436...9983) - **Exon I**; (9312...9392) - **Exon J**; (8685...9262) - **Exon K**; (8557...8635) - Exon L; (7830...7997) - Exon M; (6774...7786) - Exon N; (6648...6728) - Exon O; (6305...6598) - Exon P; (6006...6255) - Exon Q; (5732...5958) - Exon R; (4849...5076) - Exon S; (4698...4799) - Exon T; (4383...4651) - Exon U; (3336...4328) - Exon V; (2229...3094) - Exon W; (1976...2181) -

Exon X; (1635...1930) - Exon Y; (1043...1591) - Exon Z; (625...999) - Exon AA; (329...572) - Exon BB; (1...270) - Exon CC).

The LOV-1 amino acid sequence is set forth in SEQ ID NO. 4 The following table summarizes the above.

**TABLE 3 Comparison of Sequence ID No. 3 with source Cosmids†**

| EXON | SEQ ID 3 | ZK945 | F27E5 |
|---|---|---|---|
| A | 12500..12685 | | 379..564 |
| B | 12266..12451 | | 145..330 |
| C | 12085..12217 | 27844..27976 | |
| D | 11683..11823 | 27442..27582 | |
| E | 11498..11637 | 27257..27396 | |
| F | 11128..11452 | 26887..27211 | |
| G | 10268..10899 | 26027..26658 | |
| H | 10138..10216 | 25897..25975 | |
| *I | 9436..9983 | 25195..25742 | |
| *J | 9312..9392 | 25151..25071 | |
| *K | 8685..9262 | 24444..25021 | |
| L | 8557..8635 | 24316..24394 | |
| M | 7830..7997 | 23589..23756 | |
| N | 6774..7786 | 22533..23545 | |
| O | 6648..6728 | 22407..22487 | |
| P | 6305..6598 | 22064..22357 | |
| Q | 6006..6255 | 21765..22014 | |
| R | 5732..5958 | 21491..21717 | |
| S | 4849..5076 | 20608..20835 | |
| T | 4698..4799 | 20457..20558 | |
| U | 4383..4651 | 20142..20410 | |
| V | 3336..4328 | 19095..20087 | |
| **W | 2229..3094 | 17988..18853 | |
| X | 1976..2181 | 17735..17940 | |
| Y | 1635..1930 | 17394..17689 | |
| Z | 1043..1591 | 16802..17350 | |

| EXON | SEQ ID 3 | ZK945 | F27E5 |
|---|---|---|---|
| AA | 625..999 | 16384..16758 | |
| BB | 329..572 | 16088..16331 | |
| CC | 1..270 | 15760..16029 | |

*exons I, J, K at the junction of ZK945.10 and ZK945.9 (as determined by RT-PCR analysis, and not predicted by the GeneFinder program)

**the *sy582* lov-1 mutant has a 1059 bp deletion beginning in exon W at position 2267 of SEQ ID NO. 3 (18026 of the ZK945 cosmid) and ending at position 1209 of SEQ ID NO. 3 (16968 of the ZK945 cosmid).

† The GenBank accession numbers for ZK945 and F27E5 are (GenBank Accession No. Z48544) and (GenBank Accession No. Z48582), respectively.

**Exemplary knockout mutant *sy582***

A genomic deletion of *lov-1* in a PCR screen of EMS mutagenized worms was isolated. *lov-1(sy582Δ)* encodes a truncated protein lacking the polycystin/cation channel homology domain (Fig. 2d). Like *sy552, lov-1(sy582Δ)* males exhibit defects in response and Lov behaviors (Table 2), as well as low mating efficiency with *dpy-17* but not *unc-52* partners. *sy582Δ* is recessive and fails to complement *sy552.* The truncated protein produced by *lov-1(sy582Δ)* does not act as a dominant negative in contrast to the truncated protein produced by plov-1.3 (see below). This difference might be due to a dosage effect of the plov-1.3 transgene. These results confirm that the polycystin block 2/cation channel domain is essential for LOV-1 activity and indicate that *lov-1(sy582Δ)* is completely defective in LOV-1 function.

The *lov-1* (*sy582*) mutant is a 1059 bp deletion of nucleotides 18026 to 16968 of ZK945 (nucleotides 2267 to 1209 of SEQ ID NO. 3). The deletion, which begins in exon W, removes the majority of the PKD homology block 2 (a total of 308 amino acids, beginning at amino acid 2520 and ending at amino acid 2827 of the sequence set forth in SEQ ID NO. 4) and continues to read in-frame to the end of the sequence set forth in SEQ ID NO. 4. This results in a protein of 2870 amino acids with the amino acid sequence set forth in SEQ ID NO. 15.

Other mutants may be prepared by any method known to those of skill in the art, including directed mutagenesis of the gene in a selected nematode or random mutagenesis and selection for the altered male mating behavior in the lov and/or response, preferably both behaviors. Preferred regions for deletion include the exon A. Precise size of the deletion and or locations to delet can be determined empirically using standard routine methods based upon the disclosure herein, which identifies the gene and the resulting phenotype. Other mutations including insertions and point mutations that alter these behaviors are also contemplated and can be readily prepared.

**Expression patterns of *lov-1***

To elucidate the cells in which *lov-1* acts to affect male mating behaviors, the expression pattern of *lov-1*-::GFP reporter genes was examined (see Example 2 and Fig. 4). These experiments reveal regulatory regions in the *lov-1* gene. A partial translational fusion containing 2.8 kb of upstream sequence and 3.9 kb of *lov-1* (plov-1::GFP1) directs male-specific expression in male-specific sensory neurons (Fig. 2c and Fig. 4). Conversely, shorter versions of plov-1::GFP1 are not expressed in the same set of male-specific neurons nor exclusively in male-specific sensory neurons and do not act as DNs (Fig. 2c). Similar results were observed with pkd-2 mutants (see Example 2 and Fig. 4).

**Nematode pkd-2**

A search for a homolog of *LOV-1* was performed to ascertain whether nematodes possess a PKD2 ortholog. A BLAST search of the Sanger Center *C. elegans* genome data base revealed a possible *LOV-1* homolog, Y73F8A.B. This cosmid encodes a protein with 27% identity to PKD2 and possesses the coiled-coil domain of all polycystins. It is shown herein that Y73F8A.B and Y73F8A.A encode one transcript that is the *C. elegans* ortholog of human PKD2 (Fig. 2d and Fig 3). The resulting

nematode gene, designated *pkd-2*, cDNA and encoded protein are provided herein.

The *C. elegans* gene is exemplified herein. SEQ ID No. 5, which sets forth the complement of the coding strand, is provided. It contains nucleotides 1605 to 9677 of *C. elegans* cosmid Y73F8A (GenBank Accession No. AL132862), which correspond to nucleotides 1 to 8073 of SEQ ID No. 5. The sequence of the encoded protein is set forth in SEQ ID No. 6. Figure 3B shows *pkd-2* genomic structure (exons shown as boxes, introns as lines). The cDNA yk219e1 was sequenced and corresponds to the 3' end of pkd-2.

Figure 3B shows the *pkd-2* genomic structure (exons shown as boxes, introns as lines). The coding sequence in the gene set forth in SEQ ID No. 5 is produced as follows:

Complement (Join (7980...8073) - Exon 1; (7396...7585) - Exon 2; (6765...7045) - Exon 3; (5153...5283) - Exon 4; (4863...5104) - Exon 5; (3931...4158) - Exon 6; (2875...3424) - Exon 7; (1957...2208) - Exon 8; (1542...1795) - Exon 9; (367...505) - Exon 10; (1...87) - Exon 11.

As discussed above, the architecture of *LOV-1*, including a large extracellular amino terminus, Block 1, and Block 2, is similar to that of human PKD1; the architecture and sequence of *PKD-2* is similar to PKD2. Taken together, LOV-1 and PKD-2 appear to be part of a multi-component complex and pathway. Further genetic analysis of Lov behavior confirms this.

**Knockout mutation of pkd-2**

A knockout mutation can be prepared by any method known to those of skill in the art. A deletion mutant, designated *sy606* was produced (see, Examples for primers used). A 2397 bp deletion from nucleotides 8338 to 5942, starting in intron 3 and ending in intron 5, removing exons 4 and 5 (including the partial transmembrane spanning domain S1 and the polycystin motif) with the new splice in a different reading frame resulting in a stop codon (TGA) at 5736, produced a

knockout mutation. The resulting phenotype was the same as that resulting from a knockout of *lov-1*, thereby demonstrating that the two proteins are part of the same pathway that results in the observed phenotype.

The *pkd-2* (*sy606*) mutant contains a 2397 bp deletion of nucleotides 8338 to 5942 of Y73F8A (nucleotides 6734 to 4338 of SEQ ID NO. 5), starting in intron 3 and ending in intron 5, removing exons 4 and 5 (including the partial transmembrane spanning domain S1 and the polycystin motif) with the new splice in a different reading frame. This results in a stop codon (TGA) at nucleotide 5728 (nucleotide 4124 in SEQ ID NO. 5). The sequence of the protein encoded by the *pkd-2* deletion mutant (*sy606*) is set forth in SEQ ID NO. 16.

**TABLE 4**
**Comparison of Sequence ID No. 5 with source Cosmid**

| EXON | SEQ ID 5 | Y73F8A |
|---|---|---|
| 1 | 7980..8073 | 9584..9677 |
| 2 | 7396..7585 | 9000..9189 |
| 3 | 6765..7045 | 8369..8649 |
| 4 | 5153..5283 | 6757..6887 |
| 5 | 4863..5104 | 6467..6708 |
| 6 | 3931..4158 | 5535..5762 |
| 7 | 2875..3424 | 4479..5028 |
| 8 | 1957..2208 | 3561..3812 |
| 9 | 1542..1795 | 3146..3399 |
| 10 | 367..505 | 1971..2109 |
| 11 | 1..87 | 1605..1691 |

**the *sy606* pkd-2 mutant has a 2397 bp deletion of nucleotides 8338 to 5942 of Y73F8A (GenBank Accession No. AL132862; nucleotides 6734 to 4338 of SEQ ID NO. 5), starting in intron 3 and ending in intron 5, removing exons 4 and 5, with the new splice being in a different reading frame and resulting in a stop codon (TGA) at nucleotide 5728 (4124 in SEQ ID NO. 5).

Other such deletions may be similarly produced by deleting any portion that eliminates at least one of the observed phenotypic behaviors

associated with the *lov-1* and *pkd-2* pathway. Preferable targets for these deletions are those that destroy reading frame resulting in non-functional truncated proteins, deletions that eliminate transcriptional or translational control regions, deletions in the first exon or exon such that the deletion (or insertion or point mutation) eliminates or substantially attenuates activity of the encoded protein as evidenced by altered phenotype.

**The lov-1 and pkd-2 genes encode homologs of the polycystins**

It is shown herein that the *lov-1* and *pkd-2* genes and gene products are homologs of mammalian polycystins, particularly PKD1 and PKD2, respectively. As such nematodes that express these genes, and/or mutants of the genes can serve as models to study the expression of the genes, the function of these genes, to identify additional genes in the pathway, and for screening for compounds that will serve as lead compounds for treatment of PKD in mammals, particularly humans.

Neither the precise functions of the polycystins nor the molecular basis of kidney cystogenesis is known. The results provided herein show that the homologs of the polycysins act together in a pathway, that appears to be a signal transduction pathway, in sensory neurons. It has been postulated that human polycystin 1 and polycystin 2 function as an ion channel (Torres *et al.* (1998) Current Opinion in Nephrology and *Hypertension* 7:159-169). Further supporting this conlusion, are the results of others that have indicated that human PKD2 is associated with the activity of a cation channel. These results were obtained using cell-expression and electrophysiological approaches to examine the potential channel function of a protein called PCL (polycystin-like) that had been identified in the human expressed sequence-tag database by its sequence similarity with PKD2 (Chen *et al.* (1999) *Nature 401*:383-386). PCL was expressed in *Xenopus oöcytes* by microinjecting synthetic mRNA and the channel properties were studied using the two micro-electrrode voltage clamp and patch-clamp techniques. It was found that PCL is a non-

selective cation channel that is permable to sodium, potassium and calcium. It is more permeable to calcium. Thus, PCL and PKD2 may be cation-channel subunits.

Hence, as shown herein, PKD1-related proteins act as receptors that regulate the activity PKD2-related proteins. The two proteins are part of a conserved pathway that appears to be a signalling mechanism in which the translocation of ions acts as a second messenger.

**Exemplary strains**

Strains that exhibit one or more of the behaviors are provided. The strains may be prepared by mutagenizing wild-type or other strains with other desirable characteristics and selecting for those with the behavioral phenotype.

Strain PS3152 is an N2 strain with a deletion in lov-1 (*lov-1(sy582)*)

Strain PS2816 has the *lov-1(sy552)* deletion in a background with a him-5 (high incidence of males) and plg-1, which is a mutation that causes the male to use a gelatinous mating plug (which can be used to visualize mating).

Strain PS2817 is a paralyzed (unc-52) version of PS2816.

Strain PS3150 has the same deletion in a background with a him-5 (high incidence of males) and ts lethal marker (*pha-1*). A strain with a ts marker is a good recipient for transformation.

strain recipient for transformation - *pha-1* marker - , any marker can be

PS3151 is the same as PS2815 without the *plg-1*

PS3149 has a *pha-1* marker, in a *him-5* bacground and and transforemed with an extrachromosomal element containing a *lov-1::GFP1* construct and *pha-1(+)* DNA.

Anbother strain is an *him-5* strain with the *lov-1(sy582)* deletion.

PS3400 has a deletion mutation in pkd-2, it is *pkd-2(sy606)*.

PS3401 is a *him-5* strain with the *lov-1(sy582)* deletion

PS3377 is *pkd02(sy606)* in a *him-5* background.

These and other strains may be used in the assay methods described herein or in any assay that assesses the pathways and sensory functions which *lov-1* and/or *pkd-2* are involved or that can be used for identifying compounds that affect this pathway(s).

**Assays for screening compounds and for identifying mutants with observable Lov and/or response defective behavior**

Assays for identifying additional genes in the pathway, to assess the activities of proteins in the pathway, to identify regulators of gene expressions and factors involved in gene expression of genes in this pathway, and for screening for compounds that affect polycystin function are provided. Compounds that affect polycystin function in a nematode are candidates for further investigation and serve as leads for compounds that may be therapeutically useful for treating mammalian PKDs.

Identification of components of the PKD pathway will aid in understanding the etiology of the disease and permit identification of disease markers and defective genes, thereby permitting development of reagents for diagnostic tests and identification of therapeutic targets and therapeutic agents.

The assays may be adapted for high throughput methods, particularly by using multiwell plates, such as 24, 96, 384 wells or higher densities, and automating many of the steps. By using multiple wells, for example, many compounds can be screened. The results can be automated by using video or other recording means to record the behavior in each well. Viewing using such means is facilitated by visually labeling the animals, such as by introduction of reporter gene constructs that will be expressed in areas of interest, such as the vulval and tail region of the hermaphrodite, to render the animal visible to a camera. If a GFP is used, for example, the camera will be equipped with an appropriate filter to screen out all but the green glow. Other ways of making the animals visible, include, for example, use of plg-1 animals, which leave a visible gelatinous trail as they move through the agar.

Precise protocols for culturing and nematodes, producing mutants and transgenics, and for observing behaviors are well known to those of skill in the art.

**Assays using wild-type males**

**Behavioral screens**

In these assays males will be identified that exhibit abnormal behavior, particularly abnormal Lov and/or response behaviors, thereby detecting components of PKD function, signaling or regulators, or identifying compounds that are candidates for affecting PKD function, signaling or regulation. A behavioral assay is depicted in Fig. 1, and described herein.

The tests are performed by placing male nematodes on an agar surface, such as a petri dish or microtiter plate with an agar surface, that is seeded with anything, including bacteria or chemoattractants, such as NaCl, that will keep the males in a field of view. One or more mating partners, such as a hermaphrodite, is placed on the plate and the behavior is recorded, such as by direct observation, review of a video tape, or any method whereby the behavior can be recorded.

For example, observations of the behaviors can be observed using young adult hermaphrodites, such as *unc-31(e169)* hermaphrodites, on a lawn of bacteria, such as *E. coli*. The use of *unc-31* hermaphrodites, which are sluggish, makes it easier for males to keep pace with them.

For drug screening assays, the effects of a test compound are examined. The males are treated with a compound, such as by culturing them in the presence of the compound., or including the compound in the mating dish, or pretreating the males with the compound. For analysis of mutants, males from parents or grandparents that had been mutagenized with chemical and/or radiation are tested.

In either embodiment, the behavior of the males is observed by looking for one or both, preferably both, of the Lov and 'response' behaviors compared to controls, untreated males for the drug screening

assays or wild-type for the mutant assays. If behavior of the treated males differs from controls, then the compound has some activity and is selected for further analysis.

For the assays of mutants, if the behavior of the males differs from the controls, the mutation(s) are identified, such as by mapping. The mutant gene is then identified, genetically analyzed and its role in the pathway elucidated.

These methods as well as the others provided herein can be adapted for high throughput analysis, including automation, such by videotaping and image processing. For image processing the animals can be visually labeled, such as by expressing, a reporter gene, like GFP, to produce stable transgenic strain of some construct of GFP with any promoter that would direct expression with sufficient intensity or in a sufficient number of cells to visualize the behavior. For example, a glowing vulva and tail would permit visualization of the Lov and response behaviors. Suitable genes for linkage to a reporter are any that are expressed in the animal to permit such visualization. Such markers include, but are not limited to, autofluorescence of the male spicule, egl-5-gfp, and of the hermaphrodite vulval region lin-11-gfp.

Measurements can be performed by any method known to those of skill in the art (see, *e.g.*, Liu *et al.* (1995) *Neuron 14*:79-89). Briefly, measurements can be are obtained as follows: time is kept with a stopwatch or key stroke recorder on a computer to record an 'ethogram', and distances estimated by eye and confirmed from micrographs taken of the behavior. Mating behavior is sensitive to a number of variables, including the moisture level of the plates, which are not used if they are more than a week old, hermaphrodite age. Hence controls and test animals are carefully matched. At least three hermaphrodites are used per male to control for hermaphrodite specific behaviors.

**Mating efficiency assays**

As noted above, deletion of lov-1 compromises but does not abolish the ability to mate. The mutant male can mate with paralyzed or moving impaired partners. To perform these assays, wild-type males are treated with a test compound or mutagenized, and males that sire fewer cross-progeny compared to wild-type or cannot sire cross-progeny with moving partners are identified.

To detect whether the progeny are those of the males rather than the hermaphrodites, sperm defective hermaphrodites can be used. Preferably the hermaphrodites are temperature-sensitive (ts) sperm defective. Alternatively, the mating can be detected by using a visual marker, such as using short and fat (Dpy;Dumpy) hermaphrodites, or males that express a visually or otherwise detectable transgene, such as fluorescent proteins (FPs), including, but not limited to blue fluorescent proteins and green fluorescent proteins (GFPs), and looking for the transgene in progeny could have a transgene transferred into the progeny by the mating and detectable. If a FP is used as a marker, glowing offspring are detected.

Progeny can also be detected by measuring the density of the resulting culture and a ts sperm defective hermaphrodite. If there are lot of progeny, it can be inferred that the males have mated, since the hermaphrodite is sperm defective.

**Assays using mutant males**

Suppressor and enhancer genetics can be used to assign functions to genes, to assign genes to pathways, to identify the key switches in these pathways and to provide a sensitive assay to identify new genes in a pathway and lead compounds that modulate the activity of genes and/or gene products in the pathway.

**Suppressor screen** In these assays, the process starts with a *lov-1* mutant and restoration of one or both behaviors is assessed, thereby identifying compounds or mutations that restore the defect. Restoration can occur, for example, by by-passing the defective gene, such as constitutive expression of a gene further down the pathway that had previously required *lov-1* or *pkd-2* activity. Alternatively, a mutation could knock-out the activity of another gene that suppresses the activity of *lov-1* or *pkd-2*, thereby restoring the pathway. These assays will identify other genes in the pathway. These assays can also identify a compound that corrects defect in the pathway, thereby providing a promising therapeutic lead for treatment of APKD.

**Enhancer screen** In these assays, the defect is exacerbated by looking for mutations or compounds that increase the penetrance of the phenotype caused by the *lov-1* or *pkd-2* mutations for either or both of the 'response' and Lov defect. This is achieved by screening for males that cannot sire cross progeny with paralyzed hermaphrodite mating partners or by observing the behavior directly. The genes with mutations responsible for the increased penetrance that differ are identified and those that are not *lov-1* or *pkd-2* are selected. Mammalian, particularly human, homologs of the selected genes are identified, and tested to assess their role in PKD diseases, such as, for example, by screening PKD patients for alterations in the homologous (or orthologous) gene, analysis of mouse model knockout mutations, or other methods known to those of skill in the art.

**Assays for identifying the role of PKD proteins in sensory function**

As shown herein, *lov-1* and *pkd-2* are expressed in CEM neurons, indicating that they have activity in other sensory functions, such as finding a mating partner at a distance, *i.e.* sexual chemotaxis or kinesis, where the male randomly finds a hermaphrodite and then stays nearby. Hence sexual or chemoattraction assays can be used to study PKD function. To perform this assay, for example, put males that are

mutagenized or treated with a test compound on a surface containing at particular locations hermaphrodites and a control (*i.e*, males, or other hermaphrodites, or buffer), The proportion of fraction of males that choose the hermaphrodites compared to the control is scored. If the male is defective in this sensory function, it will not distinguish between males and hermaphrodites.

Other sensory functions can be assessed to identify the role, if any, of PKD genes in the functions.

**Assays that use dominant negative forms of PKD in nematodes or in other cells to identify mutations and/or compounds that inhibit or otherwise alter PKD function**

Transgenic nematodes that express a version of the LOV-1 or PK2D protein that inhibits the activity of LOV-1 and/or PKD-2 as assessed by manifestation of the altered LOV and/or response phenotypic behavior(s) are used in these assays.

As described above, a dominant negative mutation is a mutation that encodes a polypeptide that when expressed disrupts that activity of the protein encoded by the wild-type gene (see, Herskowitz (1987) *Nature 329*:219-222). A cloned gene is altered so that it encodes a mutant product that upon expression in an organism or cell containing the wild-type gene, expression of the wild-type product is inhibited or eliminated. As a result, the cell or organism is deficient in the product. The mutation is "dominant" because its phenotype is manifested in the presence of the wild-type gene, and it is "negative" in the sense that it inactivates the wild-type gene function. It is possible to do this because proteins have multiple functional sites. Hence an assay that identifies a dominant negative mutation can identify functional activities of a protein.

In this instance, the assays use transgenic nematodes that contain such a dominant negative *lov-1* or *pkd-2* transgene. In certain assays, the transgenic mutants are mutagenized, and mutants that lose a remaining activity are selected. The mutations and genes responsible for

the loss are identified. Corresponding mammalian, particularly human, genes, such as by searching databases for homologs or by probing libraries with the nematode genes, are identified.

In the compounds screening assays that employ these transgenic nematodes, compounds that interfere with a remaining activity of the lov-1 or pkd-2 gene are identified. For example, as shown herein, plov-1.3 (plov-1.3 encodes a truncated protein lacking the polycystin block 2/channel domain) has a dominant negative effect in transgenic nematodes affecting only the Lov behavior, not Response. Compounds that rescue this dominant negative effect include those that interfere with the synthesis, binding or function of the amino-terminal region of the LOV-1 protein.

Since the dominant negative effect only affects the Lov response, a stable transgenic nematode strain that expresses a dominant negative of *lov-1*, can be used to screen for compounds and mutations that further affect Response well.

**Assays based on localization and trafficking of LOV-1 and/or PKD-2 within a cell or cells**

To identify regulators and factors necessary for synthesis and transport of *LOV-1* and/or *PKD-2* proteins, strains in which LOV-1 and PKD-2 are expressed linked to a detectable label, such as a fluorescent protein, can be and have been produced. It has been shown that these proteins are expressed in the ciliated endings and in the baso-dendritic compartment of HOB, ray neurons or CEM neurons.

These strains, such as PS3149, described above, can be used to study the trafficking patterns of *LOV-1* and *PKD-2* and cellular location(s) of the proteins in the animal by looking for mutants thereof that have altered trafficking and/or altered localization of one or both of these proteins. The mutations can be mapped, genetically analyzed and the genes identified. Such genes could serve as therapeutic or diagnostic targets.

**Assays for identification of transcriptional regulators of expression of *lov-1* and/or *pkd-2***

To identify transcriptional regulators of *lov-1* or *pkd-2*, a screen for loss or alteration of expression of either gene is provided. Transgenic nematodes with a reporter gene, such as a gene encoding a FP or lacZ or other detectable product, linked to the nucleic acid encoding *lov-1* or *pkd-2* is used. The animal is mutagenized or treated with a test compound and loss of expression or reduction in expression of either gene is assessed by detecting, such as by observing under a dissecting or compound microscope or other means, including whole animal sorting, the number of cells that express the detectable marker, such as a FP.

As a control, to avoid detection or identification of non-specific effects, an unrelated gene, such as *lin-3*, linked to a reporter, is expressed in other cells in these animals. Only mutatants that exhibit changes in expression of *lov-1* or *pkd-2*, but not expression of the other gene, are selected for identification and mapping of the mutation. If expression of the other gene is affected also, then mutation is likely affecting a general process and would not be of interest.

These assays will identify regulators of and factors that affect *lov-1* and *pkd-2* expression, which regulators and factors could serve as therapeutic or diagnostic targets, or which can aid in developing an understanding of the development and progression of PKD in mammals.

**Visual screen based on clumping behavior**

Wild type adult males isolated from hermaphrodites will clump together on a plate with a lawn of bacteria. In contrast, *lov-1* and *pkd-2* mutant males do not exhibit this clumping behavior. Rather, *lov-1* and *pkd-2* mutant males are randomly dispersed in the bacterial lawn. This assay may be used for a variety of purposes, including, but not limited to, the identification of compounds that inhibit wild type male clumping behavior, compounds that restore clumping behavior to *lov-1* or *pkd-2*

mutants, and the identification of genetic supressors of *lov-1* or *pkd-2* mutants.

**Kits and diagnostic systems for performing the assays**

Kits for use in screening for use in any of the assays are provided. The kits include transgenic or wild-type nematodes or both that express either wild-type or a mutant or a transgenic form of *lov-1* and/or *pkd-2*. The nematodes may be on plates, in wells or in any form suitable for the assays. Kits containing nucleic acid encoding either of the two genes, portions thereof or vectors or plasmids containing the nucleic acids or probes based upon these sequences or reporter gene constructs containing all or portions of either or both genes and a reporter molecule are also provided. The nucleic acids may be in solution, in lyophilized or other concentrated form, or may be bound to a suitable substrate. The kits can include additional reagents for performing the assays, such reagents include any for performing any of the steps of the methods. The kits include instructions for performing the assays.

The kits may also include suitable ancillary reagents, such as the appropriate buffers and reagents. The kits may also include suitable ancillary supplies, such as microtiter plates, vials, calibrator solutions, controls, wash solutions and solid-phase supports.

The kits are typically provided in packages customarily utilized in diagnostic assays. Such packages include glass and plastic, such as polyethylene, polypropylene and polycarbonate, bottles and vials, plastic and plastic-foil laminated envelopes and the like. The packages may also include containers appropriate for use in auto analyzers. The packages typically include instructions for performing the assays.

The following examples are included for illustrative purposes only and are not intended to limit the scope of the invention.

## EXAMPLE 1

**Identification of *C. elegans* orthologs of human polycystins**

**Mating behavior and mating efficiency assays.** Males were generated by use of *him-5(e1490)* (high incidence of male) strains or by heatshock of L4 hermaphrodites (Brenner (1974) *Genetics* 77:71-94). Mating efficiency (ME) tests were performed by pairing six tester L4 males with six paralyzed *unc-52* or four actively moving *dpy-17* or N2 L4 hermaphrodites. ME is the percentage of cross progeny to total progeny (Hodgkin (1983) *Genetics 103*:43-64). Behavioral observations were done on a 0.5 cm diameter lawn of OP50 (Liu *et al. Neuron 14*:79-89). Hermaphrodites (N2 or *unc-31(e169)*) were placed on a lawn with the tester male. Behavioral phenotypes were determined by keeping time with a stopwatch and manually recording the behavioral series. In one trial, a male is observed for a minimum of 10 vulva encounters or for 10 minutes, whichever comes first. A male who does not respond to hermaphrodite contact within 10 minutes is considered response defective. Response ability reflects the percentage of males successfully responding to hermaphrodite contact. An individual male's vulva location ability was calculated as: Number of positive vulva locations/Total number of vulva encounters. Ability can vary from 100% (always locate) to 0% (never locate). Vulva location efficiency indicates the average behavior of a genotypic population. Pairwise comparisons were made using Mann-Whitney nonparametric and two-sided t tests (Instat for MacIntosh).

**Genetic screen for location of vulva (Lov mutants).** PS1395 hermaphrodites of genotype *plg-1(e2001d); him-5(e1490)* were mutagenized with EMS (Brenner (1974) *Genetics* 77:71-94). *plg-1(e2001d); him-5(e1490)* males deposit a gelatinous plug over the hermaphrodite vulva post coitum. A decrease in plugging efficiency might reflect a decrease in mating ability. An F1 clonal screen was performed by picking individual F1 progeny of mutagenized hermaphrodites to individual plates and directly observing F2 males for behavioral defects.

An F2 clonal screen was performed such that 10 F1 progeny per P0 hermaphrodite were picked to the same plate, 10 F2 hermaphrodites per F1 pool were picked to individual plates, and F3 males were observed for decreased plugging efficiency and/or location of vulva (Lov) defects. *lov-1(sy552); plg-1(e2001d); him-5* is a recessive mutation isolated in the F2 clonal screen. *lov-1(sy552)* males are response and Lov defective and also have a very low ME with *dpy-17* hermaphrodites (ME-Dpy).

**Genetic mapping of *lov-1*.** Chromosomal linkage of *lov-1(sy552)* was determined by scoring the loss of genetic markers relative to response, Lov, and ME-Dpy phenotypes, which revealed linkage between *dpy-10* and *sy552*. Further mapping was achieved via three factor crosses. From *sy552/unc-4(e120) let-25(mn25)* heterozygotes, Unc non-Let (Unc for uncoordinated, Let for lethal) recombinants were picked. As Unc males cannot mate, a test cross with *sy552* males and Unc hermaphrodites was performed to generate non-Unc *sy552/(sy552Δ)unc-25(mn25)* males. Males were scored for response, Lov, and ME-Dpy defects. 2/12 Unc non-Let recombinants segregate the *lov-1* mutant phenotype. These data placed *lov-1* between *unc-4* and *let-25*, closer to *unc-4*. Deficiency mapping indicated that *mnDf21* uncovers *sy552* whereas *eDf21* does not.

**Transformation rescue of *lov-1(sy552)* mutants.** Cosmids and plasmids (15-100 ng/μl) in the region from the right breakpoint of *eDf21* to the right breakpoint of *mnDf21* and PHA-1 (pBX, 100 ng/μl were injected into *lov-1(sy552); pha-1(e2123ts); htm-5(e1490)*. Stable lines were selected at either 19° or 25°C (Schnabel *et al.* (1990) *Science 250*:686-688). Cosmid ZK945 rescued *sy552* response and vulva location defects in four of five stable lines. A 16.9 kb HindIII fragment of ZK945 cloned into pBS(SK+) (plov1.1) containing ORFs ZK945.10 and ZK945.9 rescued *sy552* behavioral defects in 4 of 6 stable lines. A 6.7 kb HindIII-BamHI fragment of ZK945 (plov-1::GFP1) containing ORF ZK945.10 did not rescue *sy552* defects. plov-1.3 creates a frameshift at

nucleotide 17724 in ZK945 inserting a BssHII GFP fragment from plasmid pPD95.02 out of frame into the StuI site of plov-1.1 plov-1.3 fails to rescue *sy552*.

**PCR screen for genomic deletion of *lov-1*.** Approximately 315,000 haploid genomes were screened using primers designed to delete the PKD/channel domain. Primer set 1 (SEQ ID Nos. 7 and 8, respectively), the outside primers were:

JC32 5′-CTCTATTTGTGGTTCGTTGGCG-3′ and

JC36 5′-GGGAGTTTCCGTTTTCATGGGG-3′; and

internal nested primer set (SEQ ID Nos. 9 and 10, respectively) were:

JC33 5′-CTAGGACCGATGCAACAGCGAG-3′ and

JC35 5′-AACGCTGATTGGTTCAAGTGTG-3′)

are approximately 2.5 and 2.4 kb apart, respectively. One deletion allele, *lov-1(sy582Δ)* was isolated. DNA sequence analysis indicated a deletion of nucleotides 16972 to 18027 of ZK945.

**DNA-sequence analysis.** RT-PCR from *him-5(e1490)* RNA using a combination of *lov-1* primers generated overlapping cDNA clones bridging the junction between ZK945.10 and ZK945.9. Genefinder had predicted boundaries of the last exon of ZK945.10 (from position 25742 to 25174 of ZK945) and first exon of ZK945.9 (24923 to 24444). DNA sequence analysis of RT-PCR generated cDNA clones revealed three exons in the junction: one from 25742 to 25195, a second from 25151 to 25071, and a third initiating a position 25021, corresponding to exons I, J, and K, in Fig. 2b, respectively.

**PCR screen for genomic deletion of *pkd-2***

For pkd-2 the used primers (SEQ ID Nos. 11-14, respectively) were as follows:

Outside primers

LOV2.9 (Y73F8A nt 8546-8569) 5′ CCCCTCGTTTGACCATTCTATGG 3′

LOV2.10 (Y73F8A nt 8438-8457) 5′ ACGTGATCCTCTGTCGATCCAG 3′

Nested Primers

LOV2.9A(Y73F8A nt 5599-5615) 5' AGATCAAGCTGACTGCCCGTTC 3'
LOV2.10A(Y73F8A nt 5609-5631) 5'GATCCAGCGATTAGCCTTTAA CG3'/
One deletion allele, *pkd-2(sy606)* was isolated, which has a 2397 bp deletion from nucleotides 8338 to 5942 of Y73F8A (GenBank Accession No. AL132862; corresponding to nucleotides 6734 to 4338 of SEQ ID NO. 5). The deletion starts in intron 3 and ends in intron 5, removing exons 4 and 5 (including the partial transmembrane spanning domain S1 and the polycystin motif) with the new splice in a different reading frame resulting in a stop codon (TGA) at 5736, produced a knockout mutation. The resulting phenotype was the same as that resulting from a knockout of *lov-1*, thereby demonstrating that the two proteins are part of the same pathway that results in the observed phenotype.

## EXAMPLE 2

### Expression analyses of LOV-1 and PKD-2

#### Methods

GFP (see, Chalfie *et al.* (1994) *Science 263*:802-805) expression was used a marker for *lov-1* and *pkd-2* gene expression (see Figs. 3a and 4A) plov-1::GFP1 was constructed by cloning a 6.7 kb *Hind*III-*Bam*HI fragment of plov-1.1 into the vector pPD95.81, plov-1::GFP2 by cloning a *Hind*III-*Hpa*I fragment. plov-1::GFP3 and plov-1::GFP4 are *Sac*I and *Hind*III-*Hpa*I (Klenow filled-in and religated) deletions of plov-1::GFP1, respectively. plov-1::GFP5 was constructed by cloning a 15.4 kb *Hind*III-*Afe*I fragment of plov-1.1 into the *Hind*III-*Sma*I site of pPD95.79. ppkd-2.1, ppkd-2::gfp1 and ppkd-2::gfp2 were constructed by cloning PCR-amplified 8.9 kb, 2.0 kb and 5.9 kb fragments into the vectors pPD95.97, pPD95.75 and pPD95.77, respectively. Transgenic animals were observed by fluorescence microscopy Cells were identified by comparing Nomarski and fluorescent or confocal images of the same animals to determine cell-body position (Sulston *et al.* (1980) *Dev. Biol. 78*:542-576). HOB assignment was confirmed by laser ablation of precursor cells.

**lov-1 expression**

*lov-1::GFP1* is specifically expressed in male-sensory neurons, including four putative chemosensory CEM cephalic neurons, the hook neuron HOB (Fig. 4a), and the sensory ray neurons (Fig. 4b). *lov-1::GFP1* expression was first observed in a few cells during late L4 lethargus (data not shown) while strong expression peaks in the adult male. In neuronal cell bodies, GFP expression is cytoplasmic (non-nuclear) and punctate (Fig. 4a and Fig. 4b). *lov-1::GFP1* is localized at high levels in the cell body and ciliated endings of CEM (Fig. 4c), HOB, and ray neurons (Fig. 4b) but is not observed in axons. Localization of *lov-1::GFP1* to sensory endings is consistent with plasma membrane localization and strengthens the argument that *lov-1* mediates sensory perception required for mating behaviors. The temporal and spatial regulation of *lov-1* is concordant with its role in adult male mating behavior. Rays mediate response to contact with a hermaphrodite (Liu *et al. Neuron 14*:79-89), the hook mediates vulva location (Liu *et al. Neuron 14*:79-89), and the CEMs are postulated to play a role in chemosensation (Ward *et al.* (1975) *J. Comp. Neurol. 160*:313-337).

*lov-1::GFP1* expression was unaltered in *lov-1(sy552)* mutants. Expression of this fusion gene did not rescue *lov-1(sy552)* defects (Fig. 2a) and is therefore not functional. Sensory neurons and structures are normal in *lov-1(sy552)* mutants as determined by *osm-6::gfp* expression, dye filling of sensory neurons, Nomarski observation, and SEM imaging (data not shown). The defects of *lov-1(sy552)* mutants therefore cannot be attributed to abnormal development or differentiation of the response and vulva location neurons. This indicates hat *lov-1(sy552)* defects are due to defects in the function of the cells required for response and vulva location.

The Lov defect of mutations in *lov-1* is not identical to ablation of HOB, the chemosensory neuron in which *lov-1* expressed. The *lov-1* mutant and HOB-ablated males pass the vulva (Fig. 1). The *lov-1* males,

however, are capable of precisely locating the vulva, whereas HOB-ablated males resort to slow search. Therefore, the HOB neuron of *lov-1* functions, albeit in an attenuated capacity. If *lov-1(sy552)* and *lov-1(sy582Δ)* are loss of function alleles as the data suggests, then additional components are involved in Lov sensation.

Chemosensation and mechanosensation are likely involved in Lov *C elegans* sensory neurons can be polymodal: for example, by ultrastructural assignment, the ASH neuron appears to be chemosensory yet functions in both mechanosensory (nose touch) and chemosensory (osmotic avoidance) modalities (Kaplan *et al.* (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*:2227-2231). HOB might similarly be a polymodal sensory neuron. Ablation of either HOA or HOB produces identical phenotypes (Liu *et al. Neuron 14*:79-89) and HOA and HOB form multiple chemical synapses and electrical junctions (Sulston *et al.* (1980) *Dev. Biol. 78*:542-576), indicating extensive cross talk between the two hook sensory neurons. Since LOV-1 has an extensive extracellular mucin-like domain that could be involved in cell-cell or cell-matrix interaction, binding of vulva cell ligand(s) might potentially gate the LOV-1 polycystin-related channel. Another possibility is that LOV-1 could physically link the HOB sensory endings to the scherotized hook structure and couple hook deflection by the hermaphrodite vulva to intracellular voltage-activated signaling similar to hair cell mechanosensation (Hudspeth (1989) *Nature 341*:397-404) or touch response in *C. elegans* (Driscoll *et al.* in *C. elegans II* (ed. Riddle, D.I., Blumenthal, T., Meyer, B.J., and Priess, J.R.) 645-677 (Cold Spring Harbor Laboratory Press, New York, 1997).

**pkd-2 expression**

As shown herein, *C. elegans* genome contains a human PKD-2 homolog. PKD-2 possesses six membrane-spanning domains, a positively charged foruth membrane-spanning segment, a pore region, and the coiled coil domain of all polysystins. PKD-2 is localized to the same male-specific sensory neurons as LOV-1 (see, Fig. 3 and Fig. 4).

Since modifications will be apparent to those of skill in this art, it is intended that this invention be limited only by the scope of the appended claims.

SEQUENCE LISTING SUMMARY

SEQ ID No. 1 cDNA encoding human PKD1

SEQ ID No. 2 encoded human PKD1 protein

SEQ ID No. 3 sequence of a gene encoding nematode LOV-1 protein

SEQ ID No. 4 encoded nematode *LOV-1* protein

SEQ ID No. 5 sequence of a gene encoding a nematode *PKD-2* protein

SEQ ID No. 6 encoded nematode *PKD-2* protein

SEQ ID No. 7 primer for *lov-1* deletion mutant construction

SEQ ID No 8 primer for *lov-1* deletion mutant construction

SEQ ID No. 9 internal primer for *lov-1* deletion mutant construction

SEQ ID No. 10 internal primer for *lov-1* deletion mutant construction

SEQ ID No. 11 primer for *pk2-1* deletion mutant construction

SEQ ID No. 12 primer for *pk2-1* deletion mutant construction

SEQ ID No. 13 internal primer for *pk2-1* deletion mutant construction

SEQ ID No. 14 internal primer for *pk2-1* deletion mutant construction

SEQ ID No. 15 sets forth the a *LOV-1* mutant protein from *sy582*

SEQ ID No. 16 sets a *PKD-2* mutant protein from *sy606*